Applying artificial neural network theory to exploring diatom abundance at tropical Putrajaya Lake, Malaysia

Sorayya Malek\textsuperscript{a,}\textsuperscript{*}, Aishah Salleh\textsuperscript{a}, Pozi Milow\textsuperscript{a}, Mohd Sapiyan Baba\textsuperscript{b} and S.A. Sharifah\textsuperscript{c}

\textsuperscript{a}Institute of Biological Sciences, University of Malaya, Kuala Lumpur, Malaysia; \textsuperscript{b}Faculty of Computer Science and Information Technology, University of Malaya, Kuala Lumpur, Malaysia; \textsuperscript{c}Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia

(Received 12 April 2011; final version received 14 October 2011)

This article explores the relationship between diatom abundance and water quality variables in tropical Putrajaya Lake based on limnological data collected from 2001 to 2006, using supervised and unsupervised artificial neural networks (ANNs). Recurrent artificial neural network (RANN) was used for the supervised ANNs and Kohonen Self Organizing Feature Maps (SOMs) for the unsupervised ANNs. The RANN was developed for the prediction of diatom abundance using variables selected by sensitivity analysis (water temperature, pH, dissolved oxygen, and turbidity). The RANN model performance was measured using root mean squared error (19.0 cell/mL) and the r-value (0.7). SOM was used in this study for classification and clustering of diatom abundance in relation to selected water quality variables and was validated using a sensitivity curve of diatom abundance over the selected variable range generated from RANN. SOM has been employed in this study for pattern discovery of diatom abundance at Putrajaya Lake. The extracted patterns of diatom abundance in terms of propositional IF...else rules were tested and yielded an accuracy rate of 87%.

Keywords: recurrent artificial neural network; self-organizing maps; pattern discovery; tropical lakes; sensitivity analysis

Introduction

Diatoms can be important indicators of environmental conditions in lakes because they respond directly and sensitively to many physical, chemical, and biological changes in aquatic ecosystems, such as water temperature (Squires et al. 1979; Descy and Mouvet 1984), nutrient concentrations (Pringle and Bowers 1984; Pan and Stevenson 1996), and herbivory (Steinman et al. 1987; McCormick and Cairns 1994). Numerous studies have used diatoms to indicate or evaluate water quality of lakes and reservoirs. These include studies on the succession of diatom assemblages in Lake Washington (Stockner and Benson 1967), diatoms as indicators of the rate of lake acidification (Dickman et al. 1984), temperature effects on diatoms (Weckstrom et al. 1997), and on the relationships between diatoms and the environment in Spanish reservoirs (Negro and De Hoyos 2005). However, there is still a scarcity of

\textsuperscript{*}Corresponding author. Email: sorayya@um.edu.my